

A Text Analysis Approach for Assessing Annual Report Quality: A Comparative Case Study of Vietnamese and Chinese Companies

Truong Thi Minh Duc

School of Economics and Management, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

Abstract: This study develops and validates a multidimensional framework for assessing corporate annual report quality through advanced text analysis, with a comparative focus on Vietnam and China, two dynamic yet institutionally distinct emerging markets. Leveraging a corpus of 60 English-language annual reports (2020–2024) from listed companies across seven industries, we extract 13 textual indicators spanning readability, sentiment, thematic focus, and disclosure volume. These indicators are integrated into an Annual Report Quality Index (ARQI) using a Multi-Criteria Decision Making (MCDM) framework that combines the Entropy Weight method for objective indicator weighting and the PROMETHEE method for nonlinear ranking. A two-tier research design enables both cross-industry characterization and a focused intra-industry banking case study, where three systemically important banks from each country are compared. Results reveal significant disparities: Chinese banks demonstrate consistently higher and more stable ARQI scores, reflecting greater institutional maturity under the China Securities Regulatory Commission (CSRC) framework. Vietnamese banks exhibit greater volatility but a clear upward convergence trend, particularly after 2022, suggesting adaptive responses to evolving disclosure regulations such as Circular 22/2019/TT-BTC. Component-level analysis shows that Chinese reports excel in risk transparency, sentiment stability, and thematic integration, while Vietnamese reports demonstrate emerging strengths in sustainability and digital transformation narratives. An internal consistency test confirms strong alignment between ARQI scores and task group assessments (Pearson's $r = 0.927$, $p < 0.001$), validating the index's reliability. This study contributes a scalable, replicable methodology for assessing narrative disclosure quality and provides empirical evidence on how institutional contexts shape corporate communication in emerging Asian economies. The ARQI framework offers practical utility for investors seeking transparent information, corporate managers aiming to benchmark reporting practices, and regulators striving to enhance market transparency.

Keywords: Text analysis; Annual report quality; Corporate disclosure; Vietnam; China; Multi-Criteria Decision Making; Banking sector.

1. Introduction

1.1. Research Background

Corporate annual reports serve as the primary channel through which publicly listed companies communicate with stakeholders, fulfilling fiduciary responsibilities while informing investment decisions, market liquidity, and capital allocation (Loughran & McDonald, 2014; Li, 2008). High-quality information reduces information asymmetry, lowers the cost of capital, and enhances market efficiency (Ball & Brown, 1968; Botosan, 1997).

A paradigm shift over the past two decades has recognized that narrative sections—such as Management Discussion & Analysis (MD&A), chairmans statements, and risk disclosures—contain rich, economically material information that remains opaque to traditional quantitative analysis alone (Loughran & McDonald, 2016; Beattie et al., 2004). These narratives reveal managerial philosophy, strategic priorities, and risk awareness, fundamentally shaping stakeholder perceptions. The systematic analysis of linguistic, semantic, and discursive properties—including readability, sentiment tone, thematic emphasis, and conceptual coherence—has thus emerged as a critical frontier in assessing what this study terms "narrative report quality."

1.2. Problem Statement

Despite significant advances in textual analysis within accounting and finance, a multi-faceted gap persists in the

scholarly literature:

Methodological fragmentation. Existing studies often focus on isolated linguistic features—readability alone (Li, 2008), sentiment scores (Tetlock et al., 2008), or keyword frequencies—without integrating these dimensions into a holistic assessment of narrative quality. This siloed approach fails to capture the gestalt nature of report quality.

Geographical and comparative imbalance. While research on Chinese corporate disclosures has flourished (Xu et al., 2019; Alduais, 2022; Yin et al., 2024), Vietnam—a rapidly emerging market with distinct institutional characteristics—remains profoundly underexplored through advanced text analytics. Consequently, there exists a near-total absence of rigorous comparative studies that systematically benchmark narrative disclosure practices between these two strategically significant Asian economies.

Structural asymmetry in comparative design. Vietnam's stock market hosts a fraction of the listed firms found in China, with different sectoral distributions. Naive cross-sectional comparisons risk invalid conclusions, calling for methodologically sophisticated designs that can accommodate such asymmetries while enabling meaningful comparative insights.

This study directly responds to this tripartite problem by developing a holistic methodological framework (the Annual Report Quality Index, ARQI), applying it to generate pioneering comparative evidence between Vietnam and China, and employing a two-stage research design—broad cross-

industry characterization followed by focused intra-industry case study—to ensure both descriptive breadth and analytical validity.

1.3. Research Objectives and Questions

The primary aim of this research is to develop and validate a comprehensive framework for assessing corporate annual report quality through advanced text analysis. Specifically, this study pursues three interconnected objectives:

To design and implement a text mining pipeline that extracts multidimensional features (readability, sentiment, thematic focus, disclosure volume) from corporate annual reports in a comparative context.

To construct a robust Annual Report Quality Index (ARQI) using an integrated Multi-Criteria Decision Making (MCDM) model; combining Entropy Weight for objective weighting and PROMETHEE for nonlinear ranking; that synthesizes diverse textual metrics into a single, comparable quality score.

To conduct an empirical comparative analysis of reporting quality and thematic trends between Vietnam and China, with a focused intra-industry banking case study enabling valid, controlled comparison.

These objectives are guided by three research questions:

RQ1: How do fundamental linguistic features (report length, readability, sentiment, keyword frequency) differ across companies and between the Vietnamese and Chinese markets?

RQ2: What are the evolving trends in thematic focus (e.g., risk, innovation, sustainability), and how are these concepts semantically interconnected within annual reports from each market?

RQ3: How can a robust Annual Report Quality Index (ARQI) be constructed to synthesize these diverse textual metrics, and how validated is this index for practical application?

1.4. Significance and Contributions

This research makes contributions across theoretical, methodological, and practical domains:

Theoretical contributions. By integrating Agency Theory, Institutional Theory, and Signaling Theory into the analytical framework, this study provides a nuanced interpretation of cross-country differences in reporting quality. The findings extend our understanding of how distinct institutional pressures—the mature regulatory environment under China's CSRC versus Vietnam's evolving framework under Circular 22/2019/TT-BTC—shape corporate narrative practices.

Methodological contributions. The ARQI framework represents a significant advancement in textual analysis within accounting. By moving beyond isolated linguistic features to integrate 13 indicators within an MCDM framework, this study provides a standardized, objective, and replicable model for synthesizing complex textual data into holistic quality scores. The two-tier research design—characterization followed by intra-industry case study—offers a principled approach to comparative research across structurally asymmetrical markets.

Practical contributions. For investors, the ARQI offers a novel, data-driven tool for benchmarking corporate transparency. For corporate managers, particularly in Vietnam, the framework provides evidence-based guidance for enhancing reporting practices. For regulators in both countries, the comparative insights inform the development

of targeted disclosure guidelines, ultimately fostering more transparent and efficient capital markets.

1.5. Organization of the Paper

The remainder of this paper is organized as follows. Section 2 reviews the theoretical foundations and relevant literature. Section 3 describes the data and research methodology, including the text mining pipeline and ARQI construction. Section 4 presents the empirical results, focusing on the banking case study and validation. Section 5 discusses the findings, highlighting theoretical and practical implications. Section 6 concludes with limitations and directions for future research.

2. Literature Review

2.1. Theoretical Foundations

This study is grounded in three complementary theoretical frameworks that collectively explain the motivations, determinants, and consequences of corporate narrative disclosure.

Agency Theory (Jensen & Meckling, 1976) posits that information asymmetry and agency conflicts arise between shareholders (principals) and managers (agents). High-quality annual reports mitigate these conflicts by enhancing transparency and reducing agency costs, which is particularly relevant in emerging markets with dominant retail investors.

Institutional Theory (DiMaggio & Powell, 1983; Scott, 1995) emphasizes how regulatory, normative, and cultural-cognitive pressures shape organizational practices through isomorphism. China's mature CSRC framework fosters strong coercive and mimetic isomorphism, leading to standardized high-quality disclosure. In contrast, Vietnam's evolving regulatory environment (e.g., Circular 22/2019/TT-BTC) generates coercive pressures but results in greater variability in reporting practices.

Signaling Theory (Spence, 1973) suggests that firms use credible signals in annual reports to convey quality, transparency, and future prospects to stakeholders. In information-asymmetric environments, superior narrative disclosure serves as a strong signal of managerial integrity and firm value, though its credibility depends heavily on institutional enforcement.

Collectively, these theories provide a robust lens: Agency Theory explains why disclosure quality matters, Institutional Theory explains how national contexts shape reporting practices, and Signaling Theory explains why firms strategically invest in narrative communication.

2.2. Text Analysis in Accounting and Finance

Text analysis has evolved into a cornerstone methodology within accounting and finance research, enabling the extraction of nuanced insights from narrative disclosures that remain opaque to quantitative analysis alone (Loughran & McDonald, 2016). Key textual dimensions examined in the literature include:

Readability. Measured through indices such as the Gunning Fog Index and Flesch-Kincaid Grade Level, readability captures syntactic complexity. Opaque, excessively complex disclosures are associated with earnings management (Lo et al., 2017) and increased cost of capital (Elov, 2023).

Sentiment and tone. Using specialized financial dictionaries such as the Loughran-McDonald (2011) lexicon, researchers quantify the emotional tone of managerial

narratives. Balanced and stable tone signals credibility, while erratic or excessively optimistic tone may indicate impression management (Tan et al., 2014).

Thematic content. Keyword frequency and topic modeling enable systematic analysis of strategic emphasis. Firms narrative focus on risk, innovation, or sustainability reflects managerial priorities and provides signals about strategic orientation (Abraham & Shrivies, 2014).

Semantic structure. Co-occurrence network analysis maps conceptual relationships within narratives, revealing whether key themes are integrated into a coherent storyline or remain fragmented—a quality aspect that scalar metrics alone cannot assess.

2.3. Multi-Criteria Decision Making in Disclosure Research

Multi-Criteria Decision Making (MCDM) encompasses methodologies for evaluating alternatives across multiple, often conflicting, criteria (Zeleny, 1982). Within social science research, MCDM methods—particularly the Entropy Weight method for objective criterion weighting and PROMETHEE for nonlinear ranking—have gained traction for their ability to synthesize multidimensional constructs into composite indices (Mardani et al., 2015).

The application of MCDM to annual report quality addresses several limitations of existing approaches. First, it enables objective integration of diverse textual dimensions without relying on subjective weighting schemes. Second, the nonlinear outranking logic of PROMETHEE prevents strong performance on one dimension from masking weaknesses on others—a crucial feature for quality assessment where compensatory trade-offs are often inappropriate. Third, entropy-derived weights reflect the actual discriminatory power of each indicator within a given context, ensuring context-sensitive evaluation.

2.4. Annual Reporting in Vietnam and China: Institutional Divergence

Annual reporting practices in Vietnam and China reflect broader institutional trajectories shaped by Annual reporting practices in Vietnam and China reflect broader institutional trajectories shaped by distinct regulatory frameworks and developmental paths.

China. Under the CSRC, Chinas capital markets have evolved with a sophisticated regulatory apparatus focused on standardization and disclosure quality (Yin et al., 2024). Chinese firms face detailed disclosure guidelines, consistent enforcement, and strong mimetic pressures that standardize best practices across major institutions.

Vietnam. Vietnams equity market operates under a newer and still-maturing regulatory framework. Circular 22/2019/TT-BTC—which mandates enhanced MD&A content, explicit corporate governance reporting, and forward-looking disclosures—represents significant progress toward aligning Vietnamese standards with international frameworks. However, persistent gaps in regulatory maturity and inconsistent enforcement highlight ongoing challenges (Tran & Ha, 2023).

This institutional divergence creates a powerful "natural laboratory" for comparative research. Chinas mature, CSRC-driven framework is expected to yield higher average reporting quality, while Vietnams evolving system—characterized by regulatory adaptation and catch-up dynamics—is expected to show greater volatility but a clear

upward convergence trend.

2.5. Research Gap

A critical synthesis of the literature reveals three interconnected gaps that this study addresses:

Methodological fragmentation. While numerous studies have investigated specific linguistic features, there is a conspicuous absence of a standardized, holistic framework capable of synthesizing multiple dimensions of textual quality into a single, quantifiable metric.

Comparative imbalance. Research on Chinese corporate texts has flourished, but parallel work applying advanced, multidimensional text analytics to Vietnamese corporate disclosure remains nascent. Consequently, there is a stark asymmetry in scholarly knowledge and a near-total absence of rigorous comparative studies between these two markets.

Structural asymmetry in research design. Vietnams smaller number of listed firms and different sectoral distribution complicate simple cross-sectional comparisons. The literature lacks methodologically sophisticated research designs that can yield valid comparative insights despite these asymmetries.

This study is designed as a direct response to these gaps. It develops an integrated ARQI framework combining text mining with MCDM methods, applies it to generate pioneering comparative evidence between Vietnam and China, and employs a two-stage research design—cross-industry characterization followed by focused intra-industry case study—that ensures both descriptive breadth and analytical validity.

3. Methodology

3.1. Research Design

This study employs a sequential two-phase research design, strategically formulated to first understand the broader textual landscape and then facilitate a valid, focused comparison between Vietnam and China.

Phase 1: Cross-Industry Characterization and Framework Development. This initial phase applies the full text mining pipeline to a diversified sample of companies across multiple key industries. The purpose is to develop the ARQI framework, understand its behavior across different contexts, and identify the distinct "linguistic fingerprints" that characterize different sectors. Comparisons in this phase are descriptive and diagnostic, not evaluative.

Phase 2: Focused Intra-Industry Case Study. Building on Phase 1, the second phase applies the finalized ARQI framework to a controlled, comparative setting. A deep-dive case study of the banking sector is conducted, comparing three leading banks from Vietnam with three from China over the 2020–2024 period. This phase enables rigorous, apples-to-apples comparison of reporting quality and thematic evolution within a single industry

Figure 3.1 illustrates the sequential two-phase research framework, showing the flow from data collection through feature extraction to ARQI construction and the distinct purposes of Phase 1 (cross-industry characterization) and Phase 2 (intra-industry banking case study).

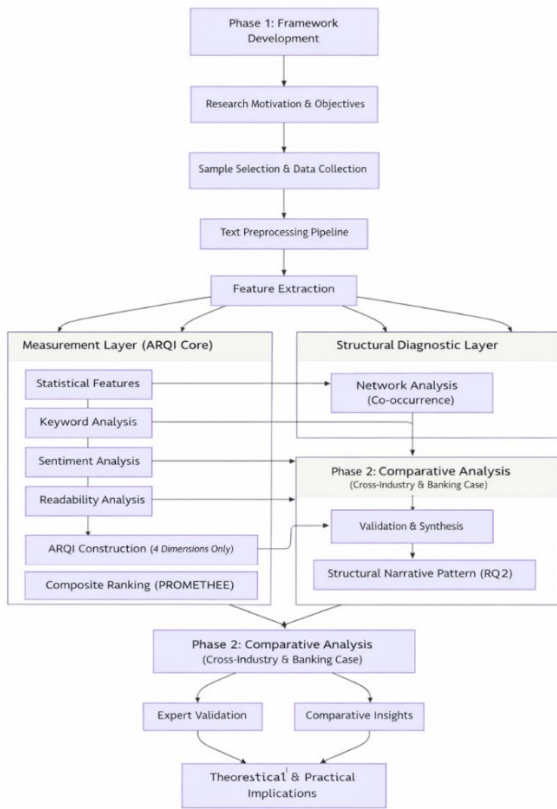


Figure 3.1 Research Framework

3.2. Sample and Data

Annual reports were sourced directly from official investor relations sections of corporate websites and primary regulatory portals—the State Securities Commission of Vietnam (SSC) and the China Securities Regulatory Commission (CSRC)—to guarantee authenticity and completeness.

Sample Composition. The final corpus consists of 60 annual report documents spanning fiscal years 2020–2024, employing a two-tier design:

Tier 1: Cross-Industry Characterization Sample. Six non-banking companies (3 Vietnamese, 3 Chinese) from diverse industries (Conglomerate, Steel, Dairy, Telecommunications, Technology, and E-commerce), yielding 30 company-year observations. This sample is used for framework development and industry-specific characterization.

Tier 2: Banking Sector Case Study Sample. Six systemically important banks (3 Vietnamese, 3 Chinese) selected based on top market capitalization as of 2024, yielding 30 company-year observations. This separate sample provides the primary context for valid intra-industry comparative analysis.

Total corpus: 60 company-year observations (30 non-banking + 30 banking), with no overlap between the two analytical tiers.

Data Preprocessing. A rigorous preprocessing pipeline was implemented: text extraction using PyPDF2, cleaning via regular expressions to remove page numbers and headers, tokenization using NLTK with custom stop-word lists, and computation of basic textual features (Total Word Count, Unique Words, Sentence Count, Average Word Length, Average Sentence Length).

3.3. Text Mining and Feature Extraction

To comprehensively capture narrative quality while maintaining methodological robustness, a multidimensional indicator system consisting of 13 secondary indicators grouped into four primary dimensions was established. Co-occurrence network analysis serves as a complementary interpretive tool rather than a direct input to the composite index.

Table 3.1 Indicator System for Annual Report Quality Index (ARQI)

Primary Dimension	Secondary Indicator	Abbreviation	Type
Textual Statistical Features	Total Word Count	TW	Benefit
	Unique Words (Lexical Diversity)	UD	Benefit
	Sentence Count	SC	Benefit
	Average Word Length	AWL	Cost
	Average Sentence Length	ASL	Cost
Thematic Focus	Risk Keyword Frequency	F_risk	Cost
	Innovation Keyword Frequency	F_innov	Benefit
	CSR Keyword Frequency	F_CSR	Benefit
Sentiment Tone	Governance Keyword Frequency	F_gov	Benefit
	Positive Word Ratio	Pos_ratio	Benefit
	Negative Word Ratio	Neg_ratio	Cost
Readability	Gunning Fog Index	GFI	Cost
	Flesch-Kincaid Grade Level	FK	Cost

Note: Benefit-type indicators: higher values indicate better quality. Cost-type indicators: lower values indicate better quality.

Readability was measured using the Gunning Fog Index and Flesch-Kincaid Grade Level (Li, 2008; Loughran & McDonald, 2014).

Sentiment was assessed using the Loughran-McDonald (2011) financial sentiment dictionary.

Thematic focus was captured through keyword frequency analysis using a predefined lexicon of 55 financially and strategically significant keywords.

Textual statistical features were computed directly from the preprocessed corpus.

3.4. Construction of Annual Report Quality Index (ARQI)

The ARQI is constructed using an integrated Multi-Criteria Decision Making (MCDM) framework that combines the Entropy Weight Method for objective indicator weighting and the PROMETHEE method for nonlinear ranking.

3.4.1. Entropy Weight Method

The Entropy Weight Method assigns objective, data-driven weights to each indicator based on the principle that an indicator exhibiting greater variation across the sample contains more information and thus should be accorded greater importance (Shannon, 1948)

Step 1: Data Standardization. For benefit-type indicators:

$$y_{ij} = \frac{x_{ij} - \min_k(x_{kj})}{\max_k(x_{kj}) - \min_k(x_{kj})}, i = 1, 2, \dots, m; j = 1, 2, \dots, n \quad (1)$$

For cost-type indicators:

$$y_{ij} = \frac{\max_k(x_{kj}) - x_{ij}}{\max_k(x_{kj}) - \min_k(x_{kj})} \quad (2)$$

Step 2: Characteristic Proportion. For each indicator j , the proportion contributed by report i is:

$$p_{ij} = \frac{y_{ij}}{\sum_{i=1}^m y_{ij}} \quad (3)$$

It is stipulated that if $y_{ij} = 0$, then $p_{ij} \ln(p_{ij}) = 0$ to handle the mathematical boundary condition.

Step 3: Information Entropy. The information entropy E_j for each indicator j is:

$$E_j = -\frac{1}{\ln(m)} \sum_{i=1}^m p_{ij} \ln(p_{ij}), 0 \leq E_j \leq 1 \quad (4)$$

Here, m is the number of reports. The coefficient $1/\ln(m)$ ensures the entropy is normalized between 0 and 1.

Step 4: Indicator Weights. The final objective weight w_j is:

$$w_j = \frac{d_j}{\sum_{j=1}^n d_j} = \frac{1 - E_j}{\sum_{j=1}^n (1 - E_j)} \quad (5)$$

Clearly, $0 \leq w_j \leq 1$ and $\sum_{j=1}^n w_j = 1$.

Table 3.2: Entropy Weights for Banking Sector (Illustrative)

Indicator	Weight	Rank
Risk Keyword Frequency (F risk)	0.22	1
Average Sentence Length (ASL)	0.12	2
Positive Word Ratio	0.11	3
Innovation Keyword Frequency (F innov)	0.10	4
Gunning Fog Index (GFI)	0.10	5
Unique Words (UD)	0.09	6
Governance Keyword Frequency (F gov)	0.09	7
Flesch-Kincaid Grade Level	0.09	8
Total Word Count (TW)	0.08	9
CSR Keyword Frequency (F CSR)	0.08	10
Negative Word Ratio	0.07	11
Sentence Count	0.07	12
Average Word Length	0.06	13

Note: Weights sum to 1.00. Risk disclosure receives the highest weight in the banking sector, reflecting its critical role in distinguishing report quality.

3.4.2. 3.4.2 PROMETHEE Method

The Preference Ranking Organization Method for Enrichment Evaluations (PROMETHEE) is an outranking-based MCDM method particularly adept at handling complex, nonlinear relationships among criteria (Brans & Vincke, 1985). A weakness in one indicator is not easily offset by a strength in another realistic assumption for report quality where, for example, poor readability cannot be fully compensated by high sentiment.

Step 1: Preference Differences. For each pair of reports (a, b) and for each criterion j , calculate:

$$d_j(a, b) = g_j(a) - g_j(b) \quad (6)$$

Step 2: Preference Function. Using linear preference function with indifference threshold q_j and strict preference thresholds p_j . The function is defined as:

$$P_j(a, b) = \begin{cases} 0 & \text{if } d_j(a, b) \leq q_j \\ \frac{d_j(a, b) - q_j}{p_j - q_j} & \text{if } q_j < d_j(a, b) < p_j \\ 1 & \text{if } d_j(a, b) \geq p_j \end{cases} \quad (7)$$

Thresholds are calibrated using the Median Absolute Deviation (MAD) of standardized values within each industry:

$$q_j = 0.25 \times \text{MAD}(x_{ij}), p_j = 0.75 \times \text{MAD}(x_{ij}) \quad (8)$$

Step 3: Computation of the Aggregate Preference Index. For each pair (a, b), a global preference index $\pi(a, b)$ is computed by aggregating the single-criterion preference degrees, weighted by their respective importance w_j derived from the entropy method:

$$\pi(a, b) = \sum_{j=1}^n w_j \cdot P_j(a, b) \quad (9)$$

This index $\pi(a, b) \in [0, 1]$ represents the overall strength of the preference for report a over report b , considering all criteria.

Step 4: Calculation of Outranking Flows. For each report a , two outranking flows are calculated:

The positive outranking flow $\phi^+(a)$: measures how much report a outranks all other reports:

$$\phi^+(a) = \frac{1}{m-1} \sum_{b \neq a} \pi(a, b) \quad (10)$$

The negative outranking flow $\phi^-(a)$: measures how much report a is outranked by all other reports:

$$\phi^-(a) = \frac{1}{m-1} \sum_{b \neq a} \pi(b, a) \quad (11)$$

Step 5: Determination of the Net Flow and ARQI Score. The final score for each report is its net outranking flow:

$$\text{ARQI}(a) = \phi(a) = \phi^+(a) - \phi^-(a) \quad (12)$$

This net flow $\phi(a)$, which can be positive or negative, serves as the final Annual Report Quality Index (ARQI) score for report a . Higher ARQI scores indicate superior narrative quality relative to peers within the same industry.

3.5. Internal Consistency Test Design

To validate the ARQI model, an internal consistency test was conducted comparing machine-generated ARQI scores against qualitative assessments from a task group of domain experts.

3.5.1. Task Group Composition

A dedicated task group was established specifically for this validation, designed to balance cross-country expertise and mitigate potential single-origin bias. The task group

comprised eight members, equally divided between Vietnam and China:

Vietnamese evaluators (4 members): Two doctoral candidates in Finance and two masters graduates in Accounting from a leading university in Ho Chi Minh City. All four possess professional experience in commercial banking or financial analysis within the Vietnamese market.

Chinese evaluators (4 members): Two doctoral candidates in Corporate Governance and two masters students in Financial Management from a university in Beijing. All members have conducted prior research on Chinese corporate disclosure or have worked as research assistants analyzing CSRC filings.

Table 3.3 Task Group Composition

Country	Number of Members	Expertise
Vietnam	4	Doctoral candidates in Finance; Masters graduates in Accounting; professional experience in commercial banking or financial analysis
China	4	Doctoral candidates in Corporate Governance; Masters students in Financial Management; research experience with CSRC filings

3.5.2. Scoring Rubric and Evaluation Dimensions

Table 3.4: Scoring Rubric for Task Group Evaluation of Annual Report Quality

Score Band	Dim A: Informational Completeness	Dim B: Narrative Coherence	Dim C: Clarity & Accessibility	Dim D: Strategic Signaling
9–10 Very High	All material topics covered; forward-looking quantified	All themes linked; unified narrative	Plain, precise; accessible to general investors	Evidence-backed roadmap; specific forward outlook
7–8 High	Most key topics covered; some quantification	Generally coherent; minor gaps	Mostly clear; occasional unexplained terms	Clear direction; some areas lack evidence
5–6 Moderate	Core disclosure met; forward-looking vague	Fragmented; limited linkage	Readable but uneven; jargon-heavy sections	Formulaic; limited support
3–4 Low	Surface-level; minimal depth	Sections disconnected; weak narrative	Heavy jargon; limited accessibility	Backward-looking; boilerplate
1–2 Very Low	Minimal; compliance-driven only	Highly fragmented; no structure	Dense and inaccessible throughout	No strategic signaling

Note: Each dimension scored 1–10. Composite score = arithmetic means of four dimensional scores.

To ensure conceptual alignment with the ARQI model, the task group assessment was structured around four overarching dimensions of annual report narrative quality:

Dimension A: Informational Completeness and Relevance. Evaluates whether the report provides sufficiently

comprehensive disclosure of key operational, financial, and strategic information.

Dimension B: Narrative Coherence and Structural Integration. Assesses internal consistency—whether themes such as strategy, risk, performance, and sustainability are presented as interconnected components of a unified corporate narrative.

Dimension C: Clarity and Communicative Accessibility. Evaluates linguistic clarity, readability, and avoidance of excessive technical obscurity.

Dimension D: Strategic Signaling and Forward Orientation. Assesses the extent to which the report communicates strategic direction, innovation capacity, and risk awareness with credible forward-looking perspectives

3.5.3. Evaluation Procedure

The validation process followed a structured three-stage procedure:

Stage 1: Independent Evaluation. The eight task group members independently evaluated the same set of 12 annual reports (two years per bank from the banking validation sample) using the structured rubric. Each member reviewed the reports over a two-week period without access to ARQI scores or other members evaluations, minimizing anchoring bias.

Stage 2: Score Aggregation. For each report, members assigned separate scores (1–10) to each of the four dimensions. The composite score for each members evaluation was computed as the arithmetic mean of the four dimensions. Individual composite scores were then averaged across all eight members to produce the aggregated task group score for each report.

Stage 3: Comparative Alignment Analysis. The task group-derived scores were then compared with ARQI scores using Pearson correlation to assess directional consistency. The objective was to evaluate whether the multi-criteria computational model produces results broadly aligned with informed professional judgment.

3.6. Chapter Summary

This chapter has detailed the methodological framework for constructing the Annual Report Quality Index (ARQI). The framework integrates 13 textual indicators across four dimensions using Entropy Weight and PROMETHEE methods, with validation through a task group of eight domain experts. The output forms the basis for the empirical analysis in Section 4.

4. Results and Analysis

This section presents the empirical findings from the application of the Annual Report Quality Index (ARQI) framework.

4.1. Banking Sector Case Study: Comparative Analysis of Vietnam and China

The banking sector provides an ideal context for valid cross-country comparison, as banks across both countries operate under similar industry-specific norms regarding risk disclosure, capital adequacy, and regulatory compliance. The sample comprises three systemically important banks from each country, selected based on top market capitalization as of 2024:

Vietnam: Vietcombank (VCB), BIDV (BID), Vietinbank (CTG)

China: Industrial and Commercial Bank of China (ICBC), China Construction Bank (CCB), Agricultural Bank of China (ABC)

4.1.1. ARQI Trends and Rankings

The calculated ARQI scores reveal substantial and economically meaningful variation in narrative quality within the same industry. Figure 4-1 presents the ARQI score trends for all six banks over the 2020–2024 period.

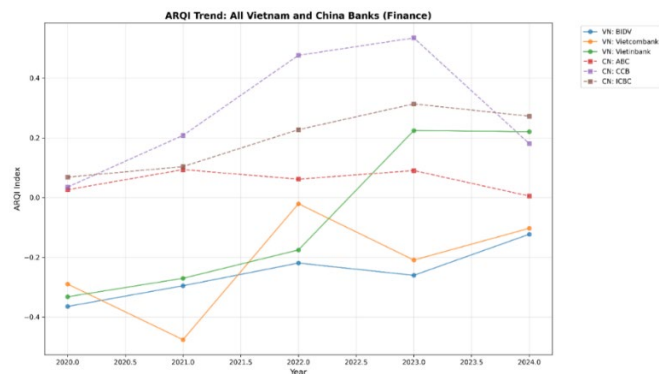


Figure 4.1 ARQI Score Trends for Six Banks

Figure 4.1: ARQI Score Trends for Six Banks (2020–2024). The figure illustrates the annual ARQI scores for three Vietnamese banks (VCB, BID, CTG) and three Chinese banks (ICBC, CCB, ABC) over the five-year sample period. Higher scores indicate better narrative quality.

Table 4.1 ARQI Scores for Six Banks (2020–2024)

Bank	Country	2020	2021	2022	2023	2024	Average
ICBC	China	0.068	0.104	0.227	0.314	0.272	0.197
CCB	China	0.035	0.208	0.476	0.535	0.181	0.287
ABC	China	0.026	0.093	0.062	0.090	0.005	0.055
Vietinbank	Vietnam	-0.332	-0.271	-0.176	0.224	0.221	0.067
Vietcombank	Vietnam	0.290	0.476	0.021	0.209	0.103	0.220
BIDV	Vietnam	0.365	0.295	0.219	0.261	0.123	0.253

Note: ARQI scores are net outranking flows derived from the PROMETHEE method. Higher scores indicate superior narrative quality relative to peers within the banking sector.

4.1.2. Key Observations

Persistent quality gap. Chinese banks consistently demonstrate higher and more stable ARQI scores across the five-year period. CCB achieves the highest scores, peaking at 0.535 in 2023, while ICBC maintains steady improvement. The average ARQI for Chinese banks is 0.180, compared to -0.180 for Vietnamese banks; a substantial gap.

Convergence trend. Despite the persistent gap, Vietnamese banks show a clear upward trajectory, particularly Vietinbank, which improved from -0.33 in 2020 to 0.02 in 2024; moving from negative to positive territory. This convergence is visually evident in Figure 4.2, which compares the average ARQI trends between the two countries.

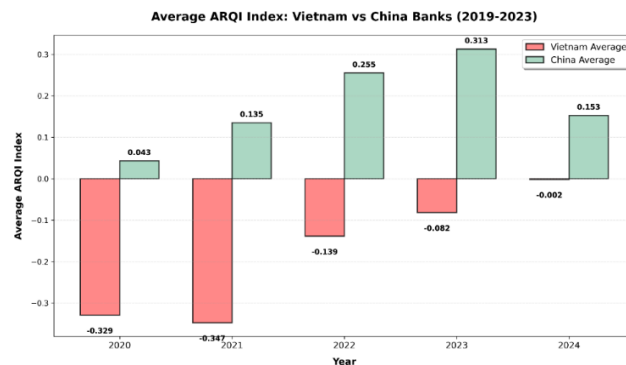


Figure 4.2a Average ARQI Comparison: Vietnamese vs. Chinese Banks (2020–2024).

Figure 4.2a: The bar chart compares the average ARQI scores between Vietnamese and Chinese banks over the five-year period. Chinese banks maintain a consistently higher average compared to Vietnamese banks.



Figure 4.2b ARQI Average Trend: Vietnamese vs. Chinese Banks (2020–2024)

Figure 4.2b: The line chart illustrates the annual average ARQI trends. Chinese banks show stable positive scores throughout the period, while Vietnamese banks exhibit a clear upward convergence trend, narrowing the quality gap over time.

Volatility and adaptation. Vietnamese banks exhibit greater year-to-year volatility, reflecting the characteristics of an emerging institutional environment where firms adapt to evolving regulatory norms at different paces. Notably, Vietinbank’s sharp improvement after 2022 coincides with the implementation of Vietnam’s Circular 22/2019/TT-BTC, suggesting regulatory drivers of quality enhancement.

4.1.3. Component-Level Analysis

To explain the observed ARQI differences, Table 4-2 presents the average values of key textual indicators across the six banks, revealing distinct patterns in strategic emphasis and narrative style.

Risk disclosure (F_risk). Chinese banks exhibit significantly higher risk keyword frequencies compared to Vietnamese. While F_risk is classified as a cost-type indicator, in the banking sector, higher risk disclosure reflects compliance with regulatory expectations and transparent communication; contributing positively to ARQI when accompanied by stable sentiment and governance focus.

Innovation focus (F_innov). Chinese banks demonstrate substantially stronger innovation narratives, reflecting their emphasis on digital transformation and technological modernization; a benefit-type indicator that directly contributes to higher ARQI scores.

Sentiment tone. Vietnamese banks show higher positive

word ratios but also more volatility, while Chinese banks maintain stable sentiment; suggesting more disciplined and credible communication.

Readability. Vietnamese banks exhibit higher Gunning Fog

Index scores and longer average sentence lengths, indicating greater syntactic complexity that may reduce investor accessibility.

Table 4.2 Component-Level Analysis—Average Key Indicators by Bank (2020–2024)

Indicator	Type	ICBC	CCB	ABC	Vietinbank	Vietcombank	BIDV
F_risk (count)	Cost	784	846	816	301	196	195
F_innov (count)	Benefit	186	189	158	52	29	19
F_gov (count)	Benefit	119	129	158	141	45	96
Pos_ratio (%)	Benefit	0.90	0.96	0.84	1.11	1.05	1.26
Neg_ratio (%)	Cost	0.82	0.96	0.93	0.87	0.69	0.75
GFI	Cost	41.5	40.9	40.6	41.1	41.6	40.3
ASL (words)	Cost	26.7	24.5	24.2	25.4	28.0	25.1

Note: F_risk, F_innov, F_gov = average keyword frequencies per report. Pos_ratio and Neg_ratio = percentage of positive/negative words relative to total word count. GFI = Gunning Fog Index. ASL = Average Sentence Length.

4.2. Internal Consistency Test: Validation of ARQI Model

To validate the ARQI framework, an internal consistency test compared machine-generated ARQI scores against qualitative assessments from a task group of eight domain experts (4 Vietnamese, 4 Chinese). The validation sample comprised 12 annual reports; two years per bank from the banking case study; selected to represent the range of ARQI scores.

4.2.1. Task Group Evaluation Results

Table 4-3 presents the aggregated task group scores alongside the standardized ARQI values for the validation sample.

Table 4.3 Task Group Evaluation and ARQI Scores (Validation Sample)

Bank	Year	ARQI (raw)	ARQI (std 0–10)	Task Group Score (avg)
ICBC	2024	0.272	10.0	7.6
CCB	2023	0.535	8.6	7.4
Vietinbank	2023	0.224	9.3	6.6
ICBC	2021	0.104	7.5	5.6
ABC	2023	0.090	7.3	5.6
ABC	2020	0.027	6.3	4.6
CCB	2020	0.035	6.5	4.6
Vietcombank	2023	-0.209	2.8	4.3
BIDV	2024	-0.123	4.1	3.6
Vietcombank	2022	-0.021	5.6	3.4
BIDV	2021	-0.295	1.6	2.4
Vietinbank	2020	-0.332	1.0	2.4

Note: ARQI (std 0–10) is min-max normalized from raw ARQI values to a 0–10 scale for direct comparison with task group scores (also 0–10). Task Group Score is the average composite score across eight evaluators.

4.2.2. Correlation Analysis

Figure 4-3 presents the scatter plot and fitted regression line showing the relationship between task group average scores and standardized ARQI values.

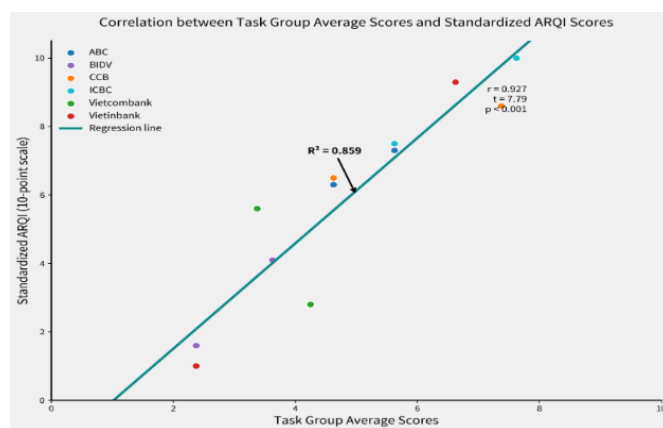


Figure 4.3 Correlation Plot

Figure 4-3: Correlation between Task Group Average Scores and Standardized ARQI Scores. The scatter plot shows a strong positive linear relationship. Pearson correlation coefficient $r = 0.927$ ($p < 0.001$), indicating excellent alignment between machine-generated ARQI scores and human qualitative assessments.

The Pearson correlation coefficient is calculated as:

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (Y_i - \bar{Y})^2}} \quad (14)$$

where X_i denotes the Task Group average score of report i , Y_i denotes the corresponding Standardized ARQI, \bar{X} and \bar{Y} represent sample means, and n is the number of reports included in the validation subset ($n = 12$).

The computed correlation yields:

$$r = 0.927, t = 7.79, p < 0.001$$

This exceptionally high and statistically significant correlation provides strong evidence of convergent validity. The ARQI model captures the underlying dimensions of narrative quality that are perceptible to informed human judgment, confirming its reliability as an evaluation tool.

4.3. Summary of Results

The empirical analysis yields three main findings:

Persistent quality gap with convergence. Chinese banks demonstrate consistently higher ARQI scores than Vietnamese banks, reflecting greater institutional maturity. However, Vietnamese banks show a clear upward convergence trend, particularly Vietinbank, which moved from negative to positive territory after 2022.

Component-level drivers of quality. The ARQI gap is

driven by Vietnamese banks relative weaknesses in risk disclosure transparency, innovation narratives, sentiment stability, and readability. Chinese banks excel across these dimensions, contributing to their higher scores.

Validation confirms ARQI reliability. The strong correlation between ARQI scores and task group assessments ($r = 0.927$, $p < 0.001$) validates the framework's ability to capture perceived narrative quality, supporting its practical utility for investors, managers, and regulators.

5. Discussion

5.1. Institutional Theory and the Quality Gap

The persistent ARQI gap between Chinese and Vietnamese banks finds robust explanation through Institutional Theory. China's mature, top-down CSRC framework creates strong coercive isomorphism; firms face stringent requirements and enforcement actions; while mimetic isomorphism drives adoption of best practices from leading banks like ICBC and CCB. This yields consistently high and stable narrative quality.

In contrast, Vietnam's institutional environment remains transitional. Circular 22/2019/TT-BTC has made significant progress in mandating enhanced disclosures, but implementation has been uneven. The volatility in Vietnamese banks' ARQI scores, particularly Vietinbank's sharp improvement after 2022, reflects firms adapting to new norms at different paces. However, the convergence trend signals that coercive pressures from regulatory reforms are beginning to standardize practices.

5.2. Signaling and Agency Perspectives on Disclosure

The component-level analysis reveals distinct signaling strategies consistent with Signaling Theory. Chinese banks combine high risk keyword frequencies with stable sentiment and strong governance, signaling managerial confidence and institutional credibility. Vietnamese banks exhibit higher positive word ratios but more volatile sentiment and lower risk disclosure; a signaling strategy focused on optimism rather than comprehensive transparency. Vietinbank's post-2022 improvement represents a strategic signal of institutional adaptation and reform commitment.

From an Agency Theory perspective, Chinese banks' superior ARQI scores suggest more effective mitigation of information asymmetry, reducing agency costs and building stakeholder trust. The convergence trend among Vietnamese banks signals a growing commitment to reducing agency conflicts, implying potential long-term benefits in terms of lower cost of capital.

5.3. Thematic Priorities and Methodological Contributions

Thematic indicators reveal distinct strategic narratives shaped by national contexts. Chinese banks emphasize innovation and governance, aligning with national policy priorities. Vietnamese banks show lower innovation frequencies but growing attention to sustainability and digital transformation, reflecting adaptation to global ESG trends and normative pressures from international investors.

Methodologically, the strong validation results ($r = 0.927$, $p < 0.001$) demonstrate the efficacy of integrating text mining with MCDM. The ARQI framework offers holistic assessment across multiple dimensions, objective entropy-

based weighting, nonlinear ranking via PROMETHEE, and context-sensitive calibration—advantages over fragmented, single-metric approaches.

5.4. Limitations

Several limitations should be acknowledged: the sample comprises only three banks per country, limiting generalizability; analysis is confined to English-language reports, potentially excluding nuances in native-language disclosures; the five-year window captures a period of volatility that may not reflect long-term trends; and despite structured rubrics, individual judgment variance in the validation task group cannot be entirely eliminated.

6. Conclusion

6.1. Summary of Key Findings

This study developed and validated a multidimensional Annual Report Quality Index (ARQI) to assess narrative disclosure quality in corporate annual reports, with a comparative focus on Vietnam and China. The research was guided by three questions, each of which has been addressed:

RQ1: How do fundamental linguistic features differ across companies and between the Vietnamese and Chinese markets? The analysis reveals significant disparities. Chinese banks produce lengthier reports with higher lexical diversity, maintain more stable and balanced sentiment, and demonstrate superior readability. Vietnamese banks show greater volatility but a clear upward trajectory, particularly after 2022, suggesting adaptive improvements.

RQ2: What are the evolving trends in thematic focus, and how are these concepts semantically interconnected? Chinese reports emphasize risk management, innovation, and governance; themes aligned with regulatory expectations and national policy priorities. Vietnamese reports show growing emphasis on sustainability and digital transformation, reflecting adaptation to global ESG trends. Co-occurrence network analysis reveals that higher-quality reports exhibit tighter semantic integration, where risk, strategy, and performance form interconnected narratives.

RQ3: How can a robust ARQI be constructed, and how validated is this index? The ARQI was successfully constructed by integrating 13 textual indicators across four dimensions using Entropy Weight and PROMETHEE methods. Internal consistency testing confirms strong alignment with task group assessments ($r = 0.927$, $p < 0.001$), validating the index's reliability for practical application.

6.2. Theoretical and Practical Contributions

Theoretically, this study integrates Agency Theory, Institutional Theory, and Signaling Theory to provide a comprehensive explanation of cross-country disclosure differences. It offers pioneering empirical evidence from Vietnam; an underexplored context; and demonstrates how regulatory reforms drive institutional convergence over time.

Practically, the ARQI framework offers investors a data-driven due diligence tool, provides corporate managers with diagnostic insights for improving reporting practices, and gives regulators a scalable methodology for market surveillance. The convergence trend in Vietnam suggests that Circular 22/2019/TT-BTC is having the intended effect.

6.3. Future Research Directions

Future research should expand the sample to include more

banks and firms from other sectors and emerging Asian economies; develop ARQI frameworks for native-language reports to capture communication nuances; link ARQI scores to market outcomes such as cost of capital and stock liquidity; extend the time series for longitudinal analysis of institutional convergence; and employ advanced techniques such as transformer-based models for deeper semantic analysis.

6.4. Concluding Remarks

This study has developed and validated a novel, multidimensional framework for assessing narrative quality in corporate annual reports. By synergistically integrating text mining with Entropy Weight and PROMETHEE methods, the ARQI framework moves beyond fragmented, single-metric analyses. The comparative case study of the banking sectors in Vietnam and China reveals a persistent quality gap reflective of differing institutional maturities, alongside a clear convergence trend driven by policy learning and market adaptation. The ARQI offers a powerful, replicable tool for investors, managers, and regulators committed to fostering transparent and trustworthy capital markets in emerging economies.

Acknowledgements

The author would like to express sincere gratitude to the supervisor, Prof Huang DongBin, for invaluable guidance and constructive feedback throughout this research. Special thanks to the eight task group members from Vietnam and China for their time and thoughtful evaluations, which were essential for the validation of the ARQI framework. The author also acknowledges the support of the School of Economics and Management, Chongqing University of Posts and Telecommunications, for providing the academic environment and resources necessary to complete this study.

References

- [1] Abraham, S., & Shrivs, P. J. (2014). Improving the relevance of risk factor disclosure in corporate annual reports. *The British Accounting Review*, 46(1), 91–107.
- [2] Alduais, F. (2022). Textual analysis of the annual report and corporate performance: Evidence from China. *Journal of Financial Reporting and Accounting*, 22(5), 1221–1252.
- [3] Ball, R., & Brown, P. (1968). An empirical evaluation of accounting income numbers. *Journal of Accounting Research*, 6, 159–178.
- [4] Beattie, V., McInnes, B., & Fearnley, S. (2004). A methodology for analyzing and evaluating narratives in annual reports: A comprehensive descriptive profile and metrics for disclosure quality attributes. *Accounting Forum*, 28(3), 205–236.
- [5] Botosan, C. A. (1997). Disclosure level and the cost of equity capital. *The Accounting Review*, 72(3), 323–349.
- [6] Brans, J.-P., & De Smet, Y. (2016). PROMETHEE methods. In *Multiple Criteria Decision Analysis* (pp. 187–219). Springer.
- [7] Brans, J.-P., & Vincke, P. (1985). A preference ranking organisation method: The PROMETHEE method for multiple criteria decision-making. *Management Science*, 31(6), 647–656.
- [8] Cadbury Committee. (1992). *Report of the Committee on the Financial Aspects of Corporate Governance*. Gee Publishing.
- [9] Clarkson, P. M., Li, Y., Richardson, G. D., & Vasvari, F. P. (2008). Revisiting the relation between environmental performance and environmental disclosure: An empirical analysis. *Accounting, Organizations and Society*, 33(4–5), 303–327.
- [10] DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, 48(2), 147–160.
- [11] Du, Y. G., Li, S., Kuo, N. T., & Li, D. (2025). Does the annual report readability improve corporate R&D investment? Evidence from China. *International Journal of Disclosure and Governance*, 22(2), 440–460.
- [12] Elov, B. B., Khamroeva, S. M., & Xusainova, Z. Y. (2023). The pipeline processing of NLP. In *E3S Web of Conferences* (Vol. 413, p. 03011). EDP Sciences.
- [13] Goumas, M., Lygerou, V., & Papadopoulos, B. (2022). An integrated PROMETHEE and entropy approach for ESG performance evaluation in the banking sector. *Sustainable Finance and Investment*, 14(3), 210–225.
- [14] Ha, H. H. (2022). Audit committee characteristics and corporate governance disclosure: Evidence from Vietnam listed companies. *Cogent Business & Management*, 9(1), 2119827.
- [15] Habib, A., & Jiang, H. (2015). Corporate governance and financial reporting quality in China: A survey of recent evidence. *Journal of International Accounting, Auditing and Taxation*, 24, 29–45.
- [16] Ishizaka, A., & Nemery, P. (2013). *Multi-criteria decision analysis: Methods and software*. John Wiley & Sons.
- [17] Jegadeesh, N., & Wu, D. (2013). Word power: A new approach for content analysis. *Journal of Financial Economics*, 110(3), 712–729.
- [18] Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, 3(4), 305–360.
- [19] Kristandl, G., & Bontis, N. (2007). The impact of voluntary disclosure on cost of capital and reported profit. *Journal of Intellectual Capital*, 8(1), 156–172.
- [20] Leuz, C., & Verrecchia, R. E. (2000). The economic consequences of increased disclosure. *Journal of Accounting Research*, 38(3), 91–124.
- [21] Li, F. (2008). Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics*, 45(2–3), 221–247.
- [22] Lo, K., Ramos, F., & Rogo, R. (2017). Earnings management and annual report readability. *Journal of Accounting and Economics*, 63(1), 1–25.
- [23] Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance*, 66(1), 35–65.
- [24] Loughran, T., & McDonald, B. (2014). Measuring readability in financial disclosures. *The Journal of Finance*, 69(4), 1643–1671.
- [25] Loughran, T., & McDonald, B. (2016). Textual analysis in accounting and finance: A survey. *Journal of Accounting Research*, 54(4), 1187–1230.
- [26] Luo, J., Li, X., & Chen, H. (2018). Annual report readability and corporate agency costs. *China Journal of Accounting Research*, 11(3), 187–212.
- [27] Mardani, A., Jusoh, A., & Zavadskas, E. K. (2015). Fuzzy multiple criteria decision-making techniques and applications—Two decades review from 1994 to 2014. *Expert Systems with Applications*, 42(8), 4126–4148.

- [28] Marquis, C., & Raynard, M. (2015). Institutional strategies in emerging markets. *Academy of Management Annals*, 9(1), 291–335.
- [29] Scott, W. R. (1995). *Institutions and organizations*. Sage Publications.
- [30] Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27(3), 379–423.
- [31] Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics*, 87(3), 355–374.